RRAM Architectures and Circuits at Nanoscale

Sangho Shin, and Sung Mo “Steve” Kang

Dept. of EE, UCSC

April 6, 2012
Outline of This Talk

I  Introduction
II  RRAM Architectures and Models
III Complementary RRAMs
IV  Summary
Memristors and Memristive Electronics

• Circuit Theoretic and Design Research
  – Memristive nonlinear dynamics
  – Compact device modeling
  – Low Power Systems Design

• Nanoelectronics- Memristive Systems
  – Ultra-dense nonvolatile memories
  – Stateful Boolean logic and resistive nanocomputing
  – Self-reconfigurable circuits and systems
  – Neuromorphic synaptic circuits and systems, memory-intensive systems for bioinformatics and other applications
Nonvolatile Resistive Memory

• Nanotechnology Enables Ultra-Dense Memory

Early memory based on p-Si/a-Si/Ag by Univ. of Michigan

Resistive Random Access Memory (RRAM)

• Key factors of importance
  – Physical Geometry
  – Process/Thermal Reliability
  – Embedding Compatibility into CMOS
  – Forming or formingless
  – 3D Stackability

• Challenges for Memory Devices
  – Number of Storage Bits/Device
  – Retention > 10 years
  – Endurance > $10^{17}$
  – Read/Write Energy < 1fJ/bit
  – Read/Write Speed < 1ns
I. Introduction

II. RRAM Architectures and Models

III. Complementary RRAMs

IV. Summary
Issues on RRAMs

- **Technology issues**
  - Reliable device process (Variability)
  - Multi-layer memory integration

- **Circuit issues**
  - Power consumption for Write/Read operations
  - Performance degradation by “Sneak currents”
  - “Data-pattern” sensitive power & read performance
How to deal with “Sneak Currents”?  

With selection devices (e.g., 1T1R or 1D1R)

Trading-off with power (\(0 \leq v_X < v_R\))

Forced bit-line voltages (e.g., TIA type)

<table>
<thead>
<tr>
<th>Performance</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing margin</td>
<td>Large</td>
<td>Narrow</td>
<td>Large</td>
</tr>
<tr>
<td>Speed</td>
<td>Fast</td>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Power</td>
<td>Low</td>
<td>Moderate ~ High</td>
<td>Low</td>
</tr>
<tr>
<td>2D capacity</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>3D stackability</td>
<td>Challenging</td>
<td>Good</td>
<td>Good</td>
</tr>
</tbody>
</table>
Sensing Behaviors of RRAMs

- Passive RRAM with Resistor Terminations
  - Sensing performance is highly dependent on the stored “Data-pattern”, due to “Sneak Currents”
  - The more low-resistance cells lead to the lower sensing margin

![Diagram of RRAM sensing](image)

Data-pattern sensitive Read performance (128x128 array)
• Analysis and Estimation of RRAM Performance \((m \times n\) array)
  - Full node analysis needs to solve for
    \[
    \left(g_S + \sum_{i \in I_U} g_{ij}\right) \cdot v_{S(k),j} = g_{k,j} \cdot v_R + \sum_{i \in I_k} \sum_{y \in I_j} g_{iy} \sum_{u \in I_U} \left(g_{iu} \cdot v_{S(k),u}\right)
    \]
    for \(j = 1, 2, \ldots, m\)
  - A worst case equivalent model has been reported in 1969.
  - Computation efficient statistical model for analysis of data- and variability-dependency
  - Sub-grouping for modeling
    • \(G_I\) : cell to be read \((k, l)\)
    • \(G_{II}\) : cells in reading row
    • \(G_{III}\) : cells in reading column
    • \(G_{IV}\) : all the others
2x2 Equivalent Memory Model

\[
\begin{bmatrix}
(g_{e.S} + \varepsilon_{e.S}) + g_{e.IV} + g_{e.II} & 0 \\
0 & g_S + g_{e.III} + g_{kI}
\end{bmatrix}
\begin{bmatrix}
\bar{v}_{S(k)} \\
v_{S(k),l}
\end{bmatrix}
= \begin{bmatrix}
g_{e.II} \\
g_{kI}
\end{bmatrix}
\cdot v_R + \frac{1}{g_{e.IV} + g_{e.III}}
\begin{bmatrix}
g_{e.IV} & g_{e.IV} & g_{e.IV} & g_{e.III} \\
g_{e.III} & g_{e.IV} & g_{e.III} & g_{e.III}
\end{bmatrix}
\begin{bmatrix}
\bar{v}_{S(k)} \\
v_{S(k),l}
\end{bmatrix}
\]

\[
V_{S(k)} = \left[G_P - C_{X(k)}\right]^{-1} \cdot G_{a(k)} \cdot v_R
\]

- 2×2 equivalent model for \(n \times m\) RRAM array
  - Each group is represented in a single equivalent admittance
  - Capable of flexible analysis and ease of simulation

\[\varepsilon_{e.S} = (m-1)\varepsilon_{II}\]
\[g_{e.S} = (m-1)g_S\]
\[g_{e.II} = (m-1)\bar{g}_{II}\]
\[g_{e.III} = (n-1)\bar{g}_{X}\]
\[g_{e.IV} = (m-1)(n-1)\bar{g}_{IV}\]

To be analyzed
2×2 Model Results

- Example for 128×128 array
  - Conditions: \( \alpha = R_{OFF}/R_{ON} = 10^3 \), \((R_{OFF} = 10 \, \text{M}\Omega, \, R_{ON} = 10 \, \text{k}\Omega)\)
  - Data-pattern dependent \( \beta (=R_{OFF}/R_S) \) is desirable

Symbols: \( n \times m \) simulation
Lines: 2×2 model

(a) Detection margin
(b) Average cell current
Array Size Dependency

- **Optimal** $R_S$ ($R_{S,\text{opt}}$)
  - Generally required to be small for large array sizes
  - Sensitive to data-pattern

- **Detection margin** ($\Delta v_S$)
  - Decreases with array size
  - Sensitive to data-pattern
  - For 1% margin, $n$ & $m < 128$
Data-Pattern Dependency

- Highly sensitive performance to data-pattern
  - $\beta_{opt}/\alpha$ is approximately linear to data probability
  - Data dependent optimum $\beta$-ratio ($\beta_{opt}=R_{OFF}/R_{S,opt}$) is desirable

- $\alpha$-ratio ($=R_{OFF}/R_{ON}$) dependency
  - Generally good for larger $\alpha$-ratios (rapidly desensitized as $\alpha > 10^2$)

![Graph showing data-pattern dependency]
Data-Dependent Adaptable $R_{S,\text{opt}}$

- Adaptable $g_S$ configuration
  - Composed of grounded replica rows and fixed resistance rows.
    - $x$ # of replica rows
    - $y$ # of $g_{\text{OFF}}$ rows
  - Expected value of $g_S$
    \[
    x = \frac{\Delta \beta_{\text{opt}}}{\alpha - 1} \approx \frac{\Delta \beta_{\text{opt}}}{\alpha}, \text{ and } y = (\beta_{\text{opt}, \min} - x)
    \]
    \[
    \Rightarrow \bar{\beta} = x(\alpha - 1)p_1 + (x + y) \approx \beta_{\text{opt}}
    \]
  - Adaptable $R_{S,\text{opt}}$ leads to both low-power and large detection margin.
Adaptable $R_{S,opt}$ Example

- 128×128 RRAM array
  - $R_{ON} = 10k\Omega$, $R_{OFF} = 10M\Omega$ ($\alpha = 10^3$)
  - $R_S$ constructions by
    - Reconnecting a part of unread RRAM ($x=5$)
    - Dedicated $R_{OFF}$ rows ($y=35$)

![Diagram showing RRAM array and calculations](image)
Worst Case vs. Self-adaptable

- **Worst case design**
  - Constant $\beta_{opt}/\alpha$ that made for the worst case data pattern
  - The worst case appears with all stored data of “1”s
  - Excess current consumption and reduced detection margin for non-worst cases

- **Self-adaptable design**
  - Adaptable $\beta_{opt}/\alpha$ that made for every data pattern
  - For every data pattern, it provides maximally available detection margin
  - Current consumption can be saved for general non-worst cases.
Performance Comparison

- **Worst case vs. self-adaptable design**
  - Self-adaptable design shows the larger $\Delta v_S$ and low $I_{cell,av}$ simultaneously, compared to the worst-case design
  - Especially for low probability cases

![Graphs showing detection margin and average cell current comparisons between worst-case and self-adaptable designs.]

Average improvement:
+ 46% for $\Delta v_S$ increment
+ 14% current reduction
**Complementary Memory Cell**

- Complementarily written two devices cell (CR-cell)
  - Provides data-dependent equivalent sense resistance \( R_{S,eq} \)

<table>
<thead>
<tr>
<th>Readout circuit configuration</th>
<th>1R-cell</th>
<th>CR-cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sense resistance</td>
<td>( R_{S,opt} = R_{OFF} / \sqrt{\alpha} )</td>
<td>&gt;</td>
</tr>
<tr>
<td>Normalized detection margin</td>
<td>( \Delta V_{S,1R} = \left( \frac{\sqrt{\alpha} - 1}{\sqrt{\alpha} + 1} \right) )</td>
<td>&gt;</td>
</tr>
<tr>
<td>Normalized average current</td>
<td>( I_{S,1R} = g_{OFF} \cdot \frac{\sqrt{\alpha}}{2} )</td>
<td>&gt;</td>
</tr>
</tbody>
</table>
Array Properties of CR-cell Array

• **Features**
  – Lower design complexity
    • With no optimization process on $R_S$
  – Data-dependent equivalent $R_S$
  – Reduced effective density
  – Doubled number of sneak paths by complementary devices

• **Data-pattern dependency**
  – Constant detection margin
    • With larger window
  – Regulated readout currents
    • With smaller values

2x2 array of CR-cell memories
**$n \times m$ Dimensional CR-cell Array**

- **Bit-line data buffers, Mode switches**
- **Word-line decoder, Mode switches**
- **Mode switches, Sense Amplifiers**
- **Reference generators**

**Diagram Details:**
- *Reading row (Group-A)*
- *Non-reading memory rows (Group-B)*
- *Complementary of Group-B (Group-C)*
- *Complementary of Group-A (Group-D)*

- **Word-line decoder, Mode switches**
- **Bit-line data buffers, Mode switches**
- **Reference generators**

- **Mode switches, Sense Amplifiers**
- **Reference generators**

- **$v_R$**
- **$v_S$**
- **$V_{REF}$**
Write-Mode Configuration

- 2-Step ‘WORD-wise’ Writing
  - Step-1: RESET for all devices of a selected word line
  - Step-2: SET for selected bit devices (other devices are halfway selected)
Read-Mode Configuration

• ‘WORD-wise’ READ
  – Comparison $V_S$ with $V_{REF}$

• Reference generation
  – Desirable to be data-dependent
  – Finding median $V_S$ by dedicated $R_{OFF}$ & $R_{ON}$ columns

\[ V_{REF} = \left( V_{REF.H} + V_{REF.L} \right) / 2 \]

– Each REF cell has complementary devices ($p_{III} = 0.5$)
– Reference buffer isolates capacitances of SAs
Performance Comparison (I)

- **Array size dependency**
  - Optimally designed 1R-cell array vs. CR-cell array
  - CR-cell memory is capable of ~4x larger size

![Graphs showing performance comparison between 1R-cell and CR-cell RRAM]

- Data-pattern independent detection performance
Performance Comparison (II)

- **Data-pattern Dependency**
  - Optimally designed 1R-cell array vs. CR-cell array
  - Lower current consumption for all cases
  - Data-independent detection performance

**Symbols**: CR-cells array
Thin lines: 1R-cells array

**Voltage Detection Window**
**Sensing Current**
Variability Effect on Read Performance

• For a 128x128 array:
  – Resistances: log-N dist. w/i $\sigma = 20\%$
  – $R_{\text{ON},0} = 10k\Omega$, $R_{\text{OFF},0} = 10M\Omega$
  – Array size=128x128
  – REF sections for every 8-bits

![Graph showing data-dependent $V_{\text{REF}}$ and $V_{S}$](image-url)
I  Introduction

II  RRAM Architectures and Models

III  Complementary RRAMs

IV  Summary
• Computation efficient RRAM model
  – Data-dependent 2x2 equivalent circuit model
  – Provides optimal design parameters for any random data

• Complementarily written RRAM cell (CR-cell) has been presented
  – Data-pattern independent sensing performance
  – Larger voltage sensing window, Regulated lower sensing current
References